Published in final edited form as:

Ultrasound Med Biol. 2018 April ; 44(4): 794–806. doi:10.1016/j.ultrasmedbio.2017.12.006.

Clinical Utility of Fetal Short-lag Spatial Coherence Imaging

Will Long[a,*], **Dongwoon Hyun**[a,b], **Kingshuk Roy Choudhury**[c], **David Bradway**[a], **Patricia McNally**[d], **Brita Boyd**[e], **Sarah Ellestad**[e], and **Gregg E. Trahey**[a,c]

[a]Department of Biomedical Engineering, Duke University, Durham, North Carolina, USA

[b]Department of Radiology, Stanford University, Stanford, California, USA

[c]Department of Radiology, Duke University Medical Center, Durham, North Carolina, USA

[d]Department of Women's and Children's Services, Duke University Hospital, Durham, North Carolina, USA

[e]Division of Maternal-Fetal Medicine, Duke University Medical Center, Durham, North Carolina, USA

Abstract

In this study, we evaluate the clinical utility of fetal short-lag spatial coherence (SLSC) imaging. Previous work has demonstrated significant improvements in image quality with fetal SLSC imaging as quantified by measurements of contrast and contrast-to-noise ratio (CNR). The objective of this study is to examine whether this improved technical efficacy is indicative of the clinical utility of SLSC imaging. Eighteen healthy volunteers in their first and second trimesters of pregnancy were scanned using a modified Siemens SC2000 clinical scanner. Raw channel data were acquired for routinely examined fetal organs and used to generate fully matched raw and post-processed harmonic B-mode and SLSC image sequences, which were subsequently optimized for dynamic range and other imaging parameters by a blinded sonographer. Optimized videos were reviewed in matched B-mode and SLSC pairs by three blinded clinicians who scored each video based on overall quality, target conspicuity, and border definition. SLSC imaging was highly favored over conventional imaging with SLSC scoring equal to (28.2 ± 10.5%) or higher than (63.9 ± 12.9%) B-mode for video pairs across all examined structures and processing conditions. Multivariate modeling shows that SLSC imaging is a significant predictor of improved image quality with $p \leq 0.002$. Expert-user scores for image quality support the application of SLSC in fetal ultrasound imaging.

Keywords

Fetal sonography; Spatial coherence; Harmonic imaging; Clutter reduction; Reader study; Image quality; Beamforming

[*]Corresponding Author: Will Long, Department of Biomedical Engineering, 1427 FCIEMAS, Box 90281, Durham, NC 27708; willie.long@duke.edu; Phone, +1(919)660-5131.

Introduction

Fetal ultrasound has been used for over 40 years in the assessment of pregnancies and remains the cornerstone of prenatal care in the United States and other developed nations. First trimester scanning provides early evaluation of congenital anomalies, pregnancy location, fetal number, gestational age, and the health of maternal pelvic structures. Second trimester scanning, usually performed between 18 and 22 weeks of gestation, provides a similar assessment with additional evaluation of more developed structures, including the heart, spine, and intracranial anatomy (ACOG, 2009; Cargill et al., 2009).

Ultrasound has demonstrated tremendous benefit to prenatal care, largely owing to the existence of ultrasound-based markers linked to fetal and maternal health. Sonographic findings in both the first and second trimesters of pregnancy have been shown to provide early and accurate detection of fetal defects ranging from structural abnormalities and infections to genetic anomalies such as aneuploidy (Sonek and Nicolaides, 2010). The application of ultrasound in prenatal care has translated to significant decreases in maternal and fetal morbidity and mortality and a reduction in the rate of preventable procedures such as postdate induction (Crowther et al., 1999; Saari-Kemppainen et al., 1990; Bennett et al., 2004).

However, the benefits of fetal ultrasound fail to extend to the full maternal population. Visibility of fetal structures is often limited by suboptimal image quality resulting from poor resolution and high levels of acoustic clutter – two dominant sources of ultrasound image degradation (Pinton et al., 2011). This is most often the case for patients of large body habitus. To counteract absorption by tissue, imaging at increased depth necessitates the use of deeper transmit foci and lower ultrasonic frequencies, which lead to poor resolution. In addition, propagation of sound through increased amounts of intervening tissue results in high levels of reverberation and focal degradation, which contribute to a temporally-stable haze commonly known as acoustic clutter (Lediju et al., 2008). Together, decreased resolution and contrast-degrading clutter severely limit the overall detectability of fetal structures in conventional ultrasound imaging.

A number of studies have demonstrated this link between maternal obesity and inadequate visualization of fetal anatomy. Hendler et al. (2004) reported increased rates of suboptimal visualization by 49.8% in cardiac and 31% in cranio-spinal fetal structures when scanning obese patients. Likewise, Dashe et al. (2009) showed that the percentage of adequate visualization of 10 key fetal structures decreases from 72% to only 30% in normal and Class 3 obese patients, respectively. Similar decreases in the visibility of fetal anatomy have also been shown for individuals who have undergone previous cesarean section or abdominal surgery (Chung et al., 2012).

Clinical scanners have integrated advanced imaging techniques to help address these limitations in ultrasound image quality. Harmonic and compound imaging are two well-established techniques which have demonstrated efficacy in the clinical setting. Harmonic imaging benefits from decreased off-axis scattering and reduced reverberation in near-field

tissue layers (Pinton et al., 2011; Tranquart et al., 1999). Compound imaging reduces the appearance of ultrasound speckle to improve the texture and uniformity of ultrasound images (Burckhardt, 1978; Berson et al., 1981; Shattuck and von Ramm, 1982). While these advanced methods have led to significant improvements in the quality of clinical ultrasound images, modern ultrasound technology has yet to fully address the major sources of degradation in fetal sonography. As a result, inadequate visualization of fetal structures and consequential non-diagnostic sonograms remain a significant problem for a subset of the population.

We have previously developed a beamforming technique, namely short-lag spatial coherence (SLSC) imaging, which forms images of the spatial coherence of backscattered echoes. This is in contrast to conventional B-mode imaging, which forms images based on echo amplitude. Previous studies have found this technique to be measurably more robust to acoustic clutter compared to conventional imaging. The benefits of clutter reduction have been linked to significant increases in contrast and contrast-to-noise ratio (CNR) in both simulation work (Lediju et al., 2011) as well as preliminary *in vivo* studies in cardiac (Bell et al., 2013), liver (Jakovljevic et al., 2013), and fetal imaging (Kakkad et al., 2014). In first trimester fetal imaging, harmonic SLSC images showed on average a 17% increase in contrast and 85% increase in CNR over corresponding harmonic B-mode images, with the most notable improvements in poor quality, high clutter imaging conditions. These improvements were attributed to both suppression of clutter as well as decreased texture noise in SLSC imaging (Kakkad et al., 2014).

Although results reveal promise in the application of SLSC imaging in fetal sonography, the clinical value of these images remains uncertain. Given that echo coherence is a unique source of contrast, the characteristics of SLSC images are distinct from those in conventional imaging (Lediju et al., 2011). Consequently, it remains unclear whether previously reported increases in contrast and CNR translate to improved imaging performance and utility from the perspective of clinical end-users.

In this study, we specifically examine the performance of fetal SLSC imaging within the clinical setting and as characterized by expert clinicians. Matched harmonic B-mode and SLSC videos of routinely examined fetal structures from eighteen subjects are acquired and optimized by a sonographer and subsequently scored by three clinicians based on clinically relevant criteria. Through comparison of expert-user scores for image quality, the work presented herein aims to extend previous results in order to directly evaluate the clinical utility and efficacy of coherence-based fetal imaging.

Materials and Methods

Short-lag spatial coherence imaging

In conventional pulse-echo ultrasound, a focused ultrasound beam is transmitted and the subsequently generated echoes are recorded by an array of transducer elements. Time delays are applied to the recorded channel data to compensate for the geometric spread of backscattered echoes and focus the echo wavefronts based on their point of origin. In B-mode imaging, these focused array signals are summed together and envelope detected to

produce an A-line, representing the amplitude of reflected echoes along the axis of the transmitted beam. This process is repeated for multiple ultrasound transmissions to collect a series of A-lines, which together form a two-dimensional (2D) B-mode image.

In short-lag spatial coherence (SLSC) imaging, rather than summing channel data to measure echo amplitude, correlations across signals of closely-spaced array elements are calculated to measure the spatial uniformity of backscattered wavefronts. An SLSC image is formed by finding estimates of the spatial correlation taken across receive elements with given spatial separation or lag. The spatial correlation can be computed as

$$\hat{R}_{\Delta}[n] = \frac{\sum_{\xi_{\Delta}} s_i[n]\, s_j[n]^*}{\sqrt{\left(\sum_{\xi_{\Delta}} |s_i[n]|^2\right)\left(\sum_{\xi_{\Delta}} |s_j[n]|^2\right)}}. \quad (1)$$

This equation estimates the normalized correlation at axial sample n of an ensemble of in-phase and quadrature (IQ) echo signal pairs with a given element separation, or lag, Δ. ξ_{Δ} denotes the complete ensemble of echo signal pairs with lag Δ, and s_i and s_j together a represent component echo signal pair within ensemble ξ_{Δ} (Hyun et al., 2017).

SLSC pixel values are found by calculating the cumulative sum of short-lag correlation estimates from lag 1 to a given maximum lag M, commonly referred to as the short-lag value:

$$V_{\mathrm{SLSC}}[n] = \sum_{\Delta x=1}^{M} \hat{R}_{\Delta x}[n] \quad (2)$$

$V_{\mathrm{SLSC}}[n]$ defines the SLSC pixel value at axial sample n with a given short-lag value M, typically defined in the range of 1 to 30% of the transmit aperture size (Lediju et al., 2011). An SLSC image is generated by calculating this value for a series of axial samples along multiple imaging lines. Unlike B-mode images, which are logarithmically compressed and displayed on a decibel scale, SLSC images are displayed on a linear scale with no compression applied. Complete descriptions of SLSC processing and imaging characteristics can be found in previous work (Lediju et al., 2011; Dahl et al., 2011, 2012; Bell et al., 2015).

Data acquisition

Eighteen volunteers were enrolled into the study following their scheduled screenings at the Duke University Hospital Fetal Diagnostic Center. Inclusion criteria were as follows: 1) adult pregnant women in their first or second trimesters of pregnancy and 2) healthy singleton pregnancies. As shown in Table 1, volunteers spanned a wide range of body habitus. The study was approved by the Duke University Institutional Review Board. Written informed consent was obtained from all volunteers prior to enrollment.

Synthetic receive sequencing and pixel-based GPU beamforming routines were implemented on a Siemens Acuson SC2000 scanner (Siemens Medical Solutions USA, Inc., Malvern, PA, USA) to enable real-time acquisition of raw channel data during live imaging. To match the clinical setup, a custom sequence was developed to perform pulse-inversion harmonic imaging using a 6C1 curvilinear transducer, transmitting at 2 MHz and receiving at 4 MHz, with manufacturer specified settings for apodization, gain, and bandwidth.

Using the modified clinical system and imaging sequence, ultrasound channel data were acquired from three different fetal structures: the bladder, stomach, and cerebral ventricles. These structures are routinely examined to assess first and second trimester fetal health (ACOG, 2009; Seeds, 1996) and, given their anechoic nature, were identified as particularly vulnerable to degradation by acoustic clutter. All scanning and data acquisition were performed by an expert sonographer with more than 25 years of experience in fetal sonography. Pulse-inversion harmonic imaging was used during the time of the exam to guide acquisition and locate structures of interest. Once an appropriate imaging window was identified, three seconds of raw pulse-inversion channel echo data were acquired and stored for offline image formation and post-processing. This process was repeated for the each target fetal structure in all eighteen volunteers. Live scanning and data acquisition were performed at 14 fps over 51 imaging lines spanning a 16° sector with 11 cm depth.

Prior to conducting this study, hydrophone measurements were made to ensure that all acoustic output levels were within regulatory limits of the United States Food and Drug Administration (FDA, 1997). Maximum measured mechanical index (MI) and intensities with the corresponding regulatory limits are listed in Table 2. For the custom sequence in this study, output levels were measured below FDA limits and thus considered safe for fetal scanning.

Sonographer image optimization

Acquired channel data were used to form fully-matched harmonic B-mode and SLSC image sequences offline. In addition to the collection of raw image sequences, a replicate dataset was generated, comprised of B-mode and SLSC image sequences equivalently processed with proprietary image post-processing algorithms supplied to our research group by the scanner manufacturer. Image post-processing is universally implemented in modern-day clinical scanners to reduce speckle and enhance the contrast and edge definition of ultrasound images after image formation (Ortiz et al., 2012).

To match the standard clinical workflow, matched image sequences were presented to the sonographer for parameter adjustment and optimization prior to being evaluated by clinical readers. To facilitate this task, a custom program was developed in Matlab (R2014b, The MathWorks, Natick, MA, USA) to provide side-by-side display of matched image pairs with controls for selecting the displayed frame and independently adjusting the dynamic range and short-lag value of each image sequence. Because the short-lag parameter is unique to SLSC imaging, reject filtering was used as a surrogate for the corresponding parameter in B-mode. Modulation of pixel reject was designed to scale the lower bound of the dynamic range to effectively simulate the fluctuations in contrast and grayscale level observed with

changing short-lag value in SLSC. This allowed for paired controls on the user interface in order to facilitate blinding of the sonographer to imaging method.

Using this program, the sonographer was instructed to select one representative frame from each image sequence and separately optimize side-by-side B-mode and SLSC images to maximize their respective target conspicuity (TC), border definition (BD), and overall quality (OQ) – criteria which are considered critical for the accurate visualization and measurement of structures examined in this study.

Prior to performing this task, instructions along with the optimization criteria in Table 3 were carefully explained to the sonographer. Five representative image sequences, consisting of data from all three structures with varying levels of image quality, were used to calibrate the sonographer to the expected range of image quality and to introduce the custom program and its controls. Following the training session, image optimization was performed on the remaining 108 matched image sequence pairs (18 subjects × 3 targets × 2 processing conditions). Final parameter settings for each image were recorded and used to generate corresponding matched harmonic B-mode and SLSC videos from the original multi-frame channel data.

Image quality assessment

A reader study was performed to assess clinician preference of sonographer-optimized harmonic B-mode and SLSC videos. Matched video pairs were reviewed in side-by-side synchronized cine loops by three blinded clinicians with between 14 and 25 years of experience in maternal-fetal medicine. Individual videos (B-mode and SLSC) were evaluated based on the same three criteria outlined in Table 3. TC, BD, and OQ of the imaged fetal structures were scored on a Likert-type scale describing subjective levels of image quality from zero to 5, with 5 corresponding to “excellent” and zero corresponding to “poor”. Readers were instructed to perform assessments while keeping in mind the clinical tasks associated with each imaged structure.

Prior to reviewing the videos, each reader was introduced to the criteria in Table 3 and guided through the scoring process using five sample video pairs generated with channel data and optimization settings from the sonographer training session. Training was intended to calibrate clinicians to the range of image qualities in the dataset and encourage full use of the scoring scale. All data and scores generated during training were excluded from analysis. Following training, scores were assigned to each video from the 108 video pairs (18 subjects × 3 targets × 2 processing conditions), making a total of 324 scores (108 × 3 criteria) per modality for each reader. Video pairs were presented sequentially, giving a total of 108 consecutive presentations, which took on average 45 minutes for each clinician to review.

Sonographer and clinician blinding

Several steps were taken to maintain blinding of the sonographer and clinicians to imaging method and post-processing. For both image optimization and assessment, matched B-mode and SLSC image sequences were uniquely randomized for all study participants, with matched B-mode and SLSC image sequences presented on random sides and pairs of matched image sequences from all subjects, structures, and processing conditions presented

in random order. Image quality assessments were futhermore performed individually by each clinician, without help or prior exposure to the study protocol or data.

Additionally, all identifiable information was removed from the presented material, including any labeling on images and user-interface controls. Captions indicating the structure of interest in each scan were included to provide participants with the minimal context necessary for image optimization and scoring. Upper portions of both SLSC and B-mode images, containing skin and connective tissue layers away from the target structure, were also masked off to reduce bias from the characteristically unique appearances of the near-field in B-mode and SLSC. This region appears bright in B-mode, but dim in SLSC due to the presence of high amplitude, incoherent reverberation clutter, as well as imperfect focusing away from the transmit focal depth.

Statistical analysis

Outcome measure scores from TC, BD, and OQ from all three readers were fit to an enhanced multivariate model to assess differences in score between matched harmonic B-mode and SLSC images. Multivariate modeling was used to examine the effect of SLSC imaging on TC, BD, and OQ, while accounting for variations in score attributed to other factors including reader, subject, processing condition, and imaging target. The multivariate model was defined as follows:

$$y_{ijklm}=\mu+b_i+r_j+\alpha_k+\chi_l+\gamma_m+\beta\mathrm{BMI}_i+\alpha\chi_{kl}+\alpha\gamma_{km}+\beta_2\mathrm{BMI}_i+\varepsilon_{ijklm}, \quad (3)$$

where y_{ijklm} is the outcome measure of interest (OQ, BD, or TC) for the i-th subject, j-th reader (j = 1, 2, 3), k-th method (k = B-mode or SLSC), l-th processing condition (l = absent or present), and m-th imaging target (m = bladder, stomach, cerebral ventricles). Variations in TC, BD, and OQ are defined in reference to a baseline mean μ, representing the score for a B-mode bladder video without post-processing from a subject with median BMI (28.14 kg/m^2). Model terms include effects of the SLSC method α_k, post-processing χ_l, target γ_m, and BMI of the i-th subject $\beta\mathrm{BMI}_i$. In addition, there are random effects to account for variation due to subject b_i and reader r_j, which are assumed to have independent zero-mean Gaussian distributions with standard deviations σ_b and σ_r, respectively. Interaction effects between SLSC and post-processing $\alpha\chi_{kl}$, SLSC and target $\alpha\gamma_{km}$, as well as SLSC and i-th subject BMI $\beta_2\mathrm{BMI}_i$ are also included to examine the effect of SLSC imaging varied by other relevant factors in the study. Note that β_2 is only defined as nonzero when the imaging method is SLSC. Finally, ε_{ijklm} accounts for random measurement error and is assumed to have an independent zero-mean Gaussian distribution with standard deviation σ_e.

Variation in outcome measure score attributed to a given term was considered statistically significant for p-values < 0.05. Model fitting was performed by restricted maximum likelihood estimation, using the *lme4* package in R (R Core Team, 2013). Q-Q plots of the estimated residuals of the model fits were used to verify normality of the random error terms.

Inter-rater agreement was evaluated using the non-parametric Spearman's correlation coefficient r_s (Gwet, 2014). Correlations were calculated to assess the consistency and monotonicity of TC, BD, and OQ scores between pairs of readers. Overall measures of agreement were obtained by averaging the correlation coefficients across all three readers.

Results and Discussion

Example images

Examples of sonographer-optimized matched images of the fetal bladder, stomach, and cerebral ventricles are shown in Figures 1 to 3. Consistent with the results of Kakkad et al. (2014), anechoic regions in SLSC exhibit decreased levels of clutter compared to the corresponding regions in B-mode. This is particularly apparent in Figure 1, which shows significant clutter suppression leading to marked improvement in the conspicuity of the fetal bladder with SLSC imaging. Similar improvements, owing to decreases in reverberation clutter near the lateral boundary of the stomach, are observed in Figure 2. In both examples, the anechoic organs appear better delineated from the surrounding tissue in SLSC, revealing vastly improved contrast and border definition. Figure 3 shows example images of the cerebral ventricles, which appear severely corrupted by both reverberation and off-axis scatter in B-mode. SLSC imaging is observed to reduce the appearance of this clutter, resulting in improved detectability and edge definition of the ventricular walls (VW) and choroid plexus (CP). Although clinical measurements are typically made on the the posterior ventricle, improvements in image quality are most notable in the anterior ventricle where clutter levels are highest.

Notable differences in texture are also observed between matched B-mode and SLSC images. As shown in Figures 1 to 3, speckle contributes to high frequency modulations in brightness, which largely degrade the appearance of B-mode images. Post-processing appears to reduce this texture to some degree, however, degradation due to the inherent speckle pattern and the resulting texture noise remain readily apparent. Consistent with the expected increases in texture SNR with SLSC imaging, coherence images are observed to have much lower texture noise compared to their B-mode counterparts. In all three examples, tissue appears more uniform and boundaries appear more contiguous, resulting in improved detectability of the targeted structures.

As shown in Figures 1 to 3, trends in both clutter suppression and reduced texture noise with SLSC imaging appear consistent between images both with and without post-processing.

Clinician preference

TC, BD, and OQ score distributions for the matched harmonic B-mode and SLSC videos are shown in Figures 4 and 5. Correlations between scores from different outcome measures range from 0.92 to 0.95. Given the strong correlation between outcome measures, we expect inferences made with respect to one outcome measure to similarly hold for others.

Comparison of clinician scores between matched harmonic B-mode and SLSC videos reveal consistent improvements in OQ with SLSC imaging for all processing conditions and structures. Across all imaged structures, readers scored SLSC higher than B-mode in 64.2%,

equally in 26.5%, and lower in 9.3% of matched videos without image post-processing. Similarly, readers rated SLSC higher than B-mode in 66.0%, equally in 24.7%, and lower in 9.3% of matched videos with image post-processing.

Clinician preference of OQ in SLSC imaging is furthermore maintained for individual structures within both raw and post-processed conditions. In videos without post-processing, readers favored SLSC over B-mode in 51.9%, 57.4%, and 83.3% of bladder, stomach, and cerebral ventricle video pairs, respectively. This was the case in 57.4%, 59.3%, and 81.5% of bladder, stomach, and cerebral ventricle video pairs, respectively, in the post-processed condition.

As shown in Figures 4 and 5, comparable results are observed for TC and BD. Across all processing conditions and structures evaluated in this study, SLSC preference is consistently higher than B-mode preference, ranging from 44.4 to 83.3%, while B-mode preference is at maximum only 16.7%.

Multivariate analysis

Clinician preference of SLSC imaging is corroborated by results of the multivariate analysis summarized in Tables 4, 5, and 6. Effects describe the independent contributions of each model term on outcome measure scores in the baseline condition, which was defined as an unprocessed B-mode bladder video from a subject with median BMI. Interaction effects, shown in italics, describe added effects resulting from the simultaneous contribution of two terms relative to the sum of their individual effects. Interactions between SLSC imaging and other model terms therefore characterize changes in the performance of SLSC imaging for different structures, processing conditions, and BMIs relative to performance of SLSC imaging in the baseline condition. In this case, positive interaction effects indicate improvements in the performance of SLSC imaging (i.e. greater increases in score going from B-mode to SLSC imaging), whereas negative interaction effects indicate the opposite.

Model fits show that imaging method is highly predictive of score for all three outcome measures. After adjusting for other factors, results indicate that SLSC imaging contributes to statistically significant increases in outcome measure scores from the baseline B-mode condition with average improvements of 22.1% in TC, 28.8% in BD, and 25.5% in OQ ($p \leq 0.002$). For B-mode images, outcome measure scores show no significant change with imaging target, post-processing, or BMI, with the exception of increased TC score in the cerebral ventricles ($p = 0.01$). OQ shows a marginally significant increase with post-processing ($p = 0.08$).

Modeled interaction effects show that baseline improvements with SLSC imaging are maintained across all conditions examined in this study. The lack of significant interaction between SLSC and the fetal stomach ($p \geq 0.37$) as well as SLSC and image post-processing ($p \geq 0.69$) indicates that the baseline increases in score with SLSC imaging are preserved for the stomach and with image post-processing. Significant positive interactions between SLSC and the cerebral ventricles ($p \leq 0.03$) as well as SLSC and subject BMI ($p = 0.03$ for BD) moreover suggest that increases in outcome measure scores from B-mode to SLSC are greater in the cerebral ventricles and for high BMI subjects relative to baseline score

increases. Together, these interaction effects along with the baseline effect of SLSC, defined with respect to an unprocessed bladder B-mode scan from a subject with median BMI, demonstrate that increases in TC, BD, and OQ with SLSC imaging are statistically significant across all fetal structures, processing conditions, and BMIs examined in this study.

Beyond the comparison of B-mode and SLSC imaging performance, interaction effects highlight interesting behavior specifically related to SLSC imaging. Positive interactions, indicating greater increases in BD score with SLSC in higher BMI subjects, are largely consistent with previous findings that show greater improvements in contrast and CNR with SLSC imaging in patients with poor B-mode image quality (Kakkad et al., 2014). We hypothesize that the clutter-suppressing effects of SLSC are more pronounced in high BMI subjects who exhibit higher levels of clutter, resulting in greater improvements in image quality going from B-mode to SLSC. It should be noted that interactions between SLSC and BMI are only shown for BD and, to a lesser extent, OQ ($p = 0.07$). Lack of interaction in TC may be attributed to its higher baseline score relative to OQ and BD. This results in less room for improvement in TC score and consequently a weaker dependence between TC and BMI. Corresponding decreases in B-mode image quality associated with increased clutter in higher BMI subjects are also not observed with statistical significance. Nonetheless, the effect of BMI is negative for all outcome measures, suggesting an average decrease in B-mode image quality with increasing subject BMI.

Interestingly, positive interactions also suggest that image quality improvements due to SLSC imaging are significantly greater in the cerebral ventricles than in either the bladder or stomach. High levels of off-axis scatter from structures such as the calvaria and ventricular walls in combination with the anechoic properties of cerebral spinal fluid make this window particularly susceptible to clutter, and consequently, more likely to benefit from clutter suppression with SLSC imaging. A similar phenomenon is observed in cardiac SLSC imaging where clutter suppression is enhanced by the presence of hyperechoic endocardial walls adjacent to anechoic blood (Bell et al., 2013). Given the significance of this interaction, we anticipate similarly improved SLSC performance in other hyperechogenic structures such as the fetal spine, femur, and extremities.

Note that the assumption of Gaussian random error terms and the validity of the multivariate models were successfully verified by Q-Q plots, all of which show remarkable adherence of the sample data to the Gaussian distribution (Appendix A). This gives confidence that the inferences drawn from the multivariate analysis in this study are accurate.

SLSC image degradation

Though the majority of video pairs in this study demonstrate significant improvement in outcome measure score with SLSC, there remains a percentage of video pairs in which conventional B-mode was favored over SLSC. These images reveal a variety of SLSC-specific sources of degradation that may be responsible for decreased SLSC imaging performance. In a subset of these cases, SLSC image sequences show lower grayscale contrast compared to their matched B-mode images. Contrast derived from differences in backscatter amplitude can be lost in SLSC in low noise environments where coherence

becomes uniformly high. This phenomenon has been well-characterized in previous studies (Dahl et al., 2011). Though target structures examined in this work were predominantly anechoic, where this effect is not as prevalent, contrast between surrounding tissue structures with smaller differences in echogenicity is often observed to decrease in SLSC, resulting in degraded image quality. In other examples, the short-lag parameter appears to be improperly adjusted, yielding either poor lateral resolution in cases where the short-lag value was set too low or noisy speckle texture in cases where the value was set too high. Though training was incorporated into this study to familiarize the sonographer with short-lag and its effects on image quality, it is expected that the optimization of this parameter was subject to a degree of user variability, especially given the complex behavior of SLSC with varying short-lag. Ultimately, further studies are needed to formally identify and characterize the sources of image degradation in SLSC due to both artifacts inherent to coherence-based contrast as well as errors in parameter selection.

Inter-rater agreement

Score distributions for each reader are presented in Figure 6. Scores are uniquely distributed with means and variances of 2.6 ± 2.1, 3.7 ± 1.4, and 1.3 ± 1.0 for Readers 1, 2, and 3, respectively. Despite differences in their overall distributions, scores from all three readers demonstrate consistent preference of SLSC imaging. OQ scores from Readers 1, 2, and 3 show SLSC preference for 79.6%, 54.6%, and 61.6% of video pairs, respectively, with similar results for BD and TC.

Spearman's correlations, measuring the monotonicity of scores between readers, are reported in Table 7. Average correlation values across all three readers are 0.59 ± 0.11, 0.65 ± 0.07, and 0.66 ± 0.08 for TC, BD, and OQ, respectively, indicating moderate to strong inter-rater agreement. Note that correlations between Reader 3 and the other readers are consistently lower, largely due to the lower variance of scores from Reader 3 relative to the variance of scores from either Reader 1 or 2.

Study limitations

While the results of this clinical study are promising, there are limitations that should be considered in future work. Despite efforts to match and randomize the presentation of image pairs, the characteristic appearances of B-mode and SLSC limited our ability to maintain a fully blinded study. Future studies should employ metrics derived from anatomical measurements or disease state rather than direct assessment of image quality to reduce the effect of reader subjectivity. Additionally, results were representative of only a small subset of clinicians and subjects from the same institution. Multi-center clinical trials are required to generalize these findings to larger, more diverse groups of patients and clinicians. Such studies would also provide the context to assess the performance of SLSC in other fetal structures and in both healthy and abnormal pregnancies. Likewise, image acquisition and optimization were performed by a single sonographer. Additional work is needed to examine the consistency of SLSC parameter optimization across multiple sonographers, particularly with regards to the short-lag value. The average short-lag value selected by the sonographer in this study was $M = 5.3 \pm 2.9\%$, indicating that subjective image quality was maximized at relatively low short-lag values associated with lower texture noise and decreased resolution.

Studies incorporating multiple sonographers would help inform the selection of this novel parameter and assess variability in the interpretation of SLSC images with varying short-lag. In addition, as the experimental platform used in this study was developed for a single transducer, probe bandwidth limitations restricted our ability to assess imaging performance at higher frequencies commonly used in fetal imaging. However, given that axial resolution scales equivalently for both B-mode and SLSC and that clutter suppression with SLSC is largely independent of frequency (Lediju et al., 2011), the image quality improvements with SLSC observed in this study are expected to hold for the full range of frequencies used clinically. Finally, image post-processing filters used in this study were designed by the manufacturer specifically for B-mode imaging. Further investigation is needed to develop post-processing schemes optimized for the lag-dependent contrast and texture of SLSC. Despite this drawback, post-processed SLSC images in this study still demonstrated significantly improved image quality over their identically processed B-mode counterparts.

Conclusions

We have evaluated and demonstrated the clinical utility of SLSC imaging in a blinded clinical study examining clinician preference of fetal images generated using SLSC and conventional beamforming. This was performed for sonographer-optimized matched harmonic B-mode and SLSC images of the fetal bladder, stomach, and cerebral ventricles acquired from 18 volunteers in their first and second trimesters of pregnancy. Clinician scores for target conspicuity, border definition, and overall quality revealed significant improvements in subjective image quality with SLSC relative to conventional B-mode imaging with indications that SLSC is particularly well-suited for high BMI subjects and imaging of the cerebral ventricles. As demonstrated by this study, fetal SLSC imaging is a clinically useful technique, which is preferred over conventional imaging by experienced clinical end-users for visualizing key fetal structures. Overall, integration of SLSC imaging into the standard clinical workflow shows promise in improving the visibility of clinically relevant fetal structures and consequently in reducing the likelihood of non-diagnostic fetal exams.

Acknowledgments

The authors would like to thank Chad A. Grotegut, M.D., for his participation in the clinical study as well as Siemens Medical Solutions USA, Ultrasound Division, for their in-kind and technical support. This research is supported by the National Institutes of Health grants RO1-EB017711 and T32-EB001040.

Appendix A. Model Validity

The assumption of Gaussian random error terms was evaluated by plotting the quantiles of the estimated model fit residuals for each outcome measure against the quantiles of the standard normal distribution. As shown in Figure 7, Q-Q plots for all three outcome measures show remarkable adherence to the Gaussian distribution, supporting the validity of model results.

References

ACOG. Practice bulletin no. 101: Ultrasonography in pregnancy. Obstetrics and gynecology. 2009; 113:451. [PubMed: 19155920]

Bell MAL, Dahl JJ, Trahey GE. Resolution and brightness characteristics of short-lag spatial coherence (slsc) images. IEEE transactions on ultrasonics, ferroelectrics, and frequency control. 2015; 62:1265–1276.

Bell MAL, Goswami R, Kisslo JA, Dahl JJ, Trahey GE. Short-lag spatial coherence imaging of cardiac ultrasound data: Initial clinical results. Ultrasound in medicine & biology. 2013; 39:1861–1874. [PubMed: 23932276]

Bennett KA, Crane JM, O'Shea P, Lacelle J, Hutchens D, Copel JA. First trimester ultrasound screening is effective in reducing postterm labor induction rates: a randomized controlled trial. American journal of obstetrics and gynecology. 2004; 190:1077–1081. [PubMed: 15118645]

Berson M, Roncin A, Pourcelot L. Compound scanning with an electrically steered beam. Ultrasonic imaging. 1981; 3:303–308.

Burckhardt CB. Speckle in ultrasound b-mode scans. IEEE Transactions on sonics and ultrasonics. 1978; 25:1–6.

Cargill Y, Morin L, Bly S, Butt K, Denis N, Gagnon R, Hietala-Coyle MA, Lim K, Ouellet A, Racicot MH, et al. Content of a complete routine second trimester obstetrical ultrasound examination and report. Journal of Obstetrics and Gynaecology Canada. 2009; 31:272–275. [PubMed: 19416575]

Chung JH, Pelayo R, Hatfield TJ, Speir VJ, Wu J, Caughey AB. Limitations of the fetal anatomic survey via ultrasound in the obese obstetrical population. The Journal of Maternal-Fetal & Neonatal Medicine. 2012; 25:1945–1949. [PubMed: 22384816]

Crowther CA, Kornman L, O'Callaghan S, George K, Furness M, Willson K. Is an ultrasound assessment of gestational age at the first antenatal visit of value? a randomised clinical trial. BJOG: An International Journal of Obstetrics & Gynaecology. 1999; 106:1273–1279.

Dahl JJ, Hyun D, Lediju M, Trahey GE. Lesion detectability in diagnostic ultrasound with short-lag spatial coherence imaging. Ultrasonic imaging. 2011; 33:119–133. [PubMed: 21710827]

Dahl JJ, Jakovljevic M, Pinton GF, Trahey GE. Harmonic spatial coherence imaging: An ultrasonic imaging method based on backscatter coherence. IEEE transactions on ultrasonics, ferroelectrics, and frequency control. 2012:59.

Dashe JS, McIntire DD, Twickler DM. Maternal obesity limits the ultrasound evaluation of fetal anatomy. Journal of ultrasound in medicine. 2009; 28:1025–1030. [PubMed: 19643785]

FDA. Information for manufacturers seeking marketing clearance of diagnostic ultrasound systems and transducers. Rockville, MD: Center for Devices and Radiological Health, US Food and Drug Administration; 1997.

Gwet KL. Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters. Advanced Analytics, LLC. 2014

Hendler I, Blackwell S, Bujold E, Treadwell M, Wolfe H, Sokol R, Sorokin Y. The impact of maternal obesity on midtrimester sonographic visualization of fetal cardiac and craniospinal structures. International journal of obesity. 2004; 28:1607–1611. [PubMed: 15303105]

Hyun D, Crowley ALC, Dahl JJ. Efficient strategies for estimating the spatial coherence of backscatter. IEEE transactions on ultrasonics, ferroelectrics, and frequency control. 2017; 64:500–513.

Jakovljevic M, Trahey GE, Nelson RC, Dahl JJ. In vivo application of short-lag spatial coherence imaging in human liver. Ultrasound in medicine & biology. 2013; 39:534–542. [PubMed: 23347642]

Kakkad V, Dahl J, Ellestad S, Trahey G. In vivo application of short-lag spatial coherence and harmonic spatial coherence imaging in fetal ultrasound. Ultrasonic imaging. 2014 0161734614547281.

Lediju MA, Pihl MJ, Dahl JJ, Trahey GE. Quantitative assessment of the magnitude, impact and spatial extent of ultrasonic clutter. Ultrasonic imaging. 2008; 30:151–168. [PubMed: 19149461]

Lediju MA, Trahey GE, Byram BC, Dahl JJ. Short-lag spatial coherence of backscattered echoes: Imaging characteristics. IEEE transactions on ultrasonics, ferroelectrics, and frequency control. 2011; 58:1377–1388.

Ortiz SHC, Chiu T, Fox MD. Ultrasound image enhancement: A review. Biomedical Signal Processing and Control. 2012; 7:419–428.

Pinton GF, Trahey GE, Dahl JJ. Sources of image degradation in fundamental and harmonic ultrasound imaging using nonlinear, full-wave simulations. IEEE transactions on ultrasonics, ferroelectrics, and frequency control. 2011; 58:754–765.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2013. URL http://www.R-project.org/

Saari-Kemppainen A, Karjalainen O, Ylostalo P, Heinonen OP. Ultrasound screening and perinatal mortality: controlled trial of systematic one-stage screening in pregnancy. The Lancet. 1990; 336:387–391.

Seeds JW. The routine or screening obstetrical ultrasound examination. Clinical obstetrics and gynecology. 1996; 39:814–830. [PubMed: 8934033]

Shattuck DP, von Ramm OT. Compound scanning with a phased array. Ultrasonic Imaging. 1982; 4:93–107. [PubMed: 7201697]

Sonek J, Nicolaides K. Additional first-trimester ultrasound markers. Clinics in laboratory medicine. 2010; 30:573–592. [PubMed: 20638573]

Tranquart F, Grenier N, Eder V, Pourcelot L. Clinical use of ultrasound tissue harmonic imaging. Ultrasound in medicine & biology. 1999; 25:889–894. [PubMed: 10461715]

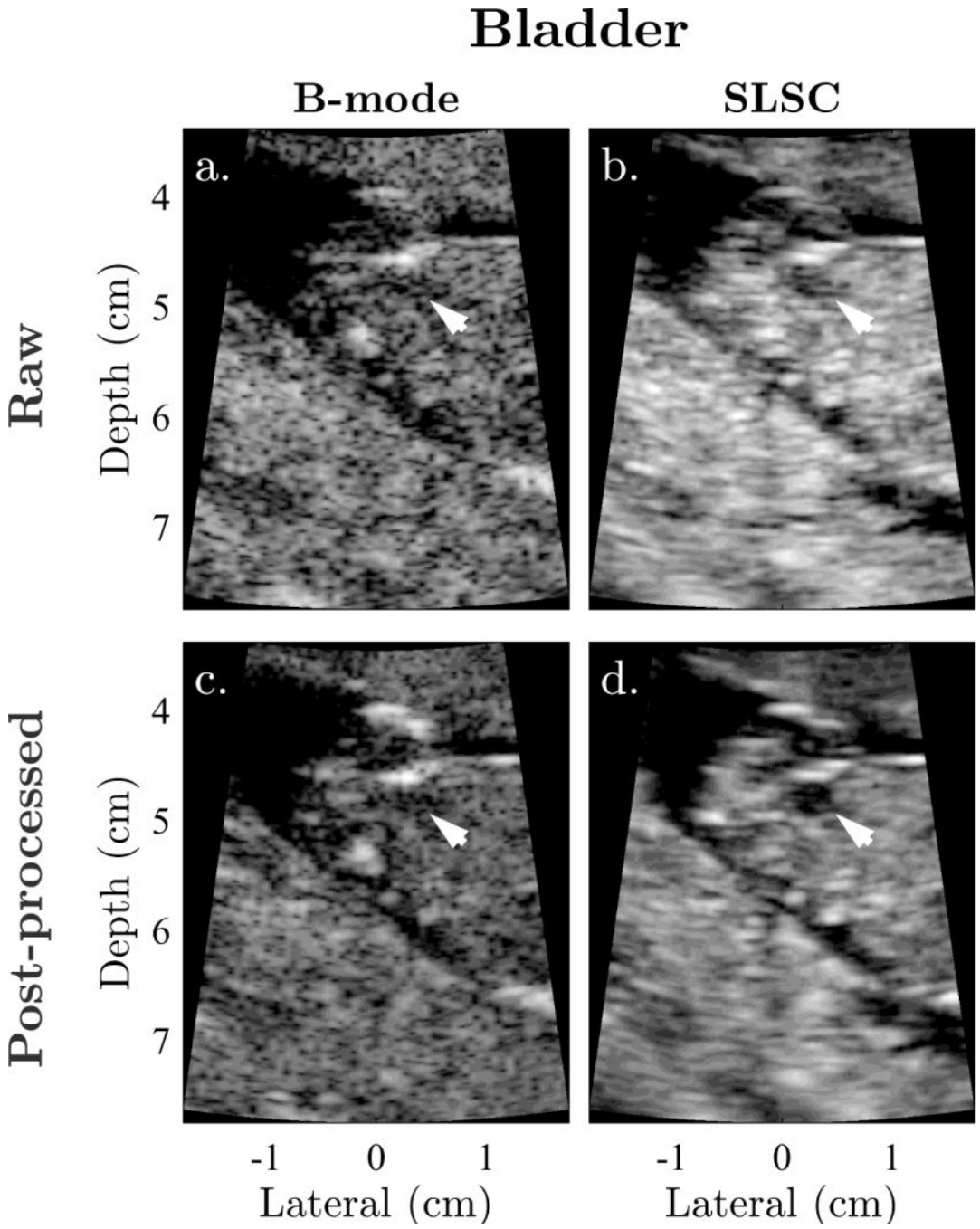


Figure 1.

Example matched (a) raw B-mode, (b) raw SLSC, (c) post-processed B-mode, and (d) post-processed SLSC images of the fetal bladder shown with sonographer-optimized parameters for dynamic range and short-lag. Arrows indicate the location of the bladder. The anechoic organ is significantly degraded by clutter in B-mode, but readily visualized with high contrast in both raw and post-processed SLSC images.

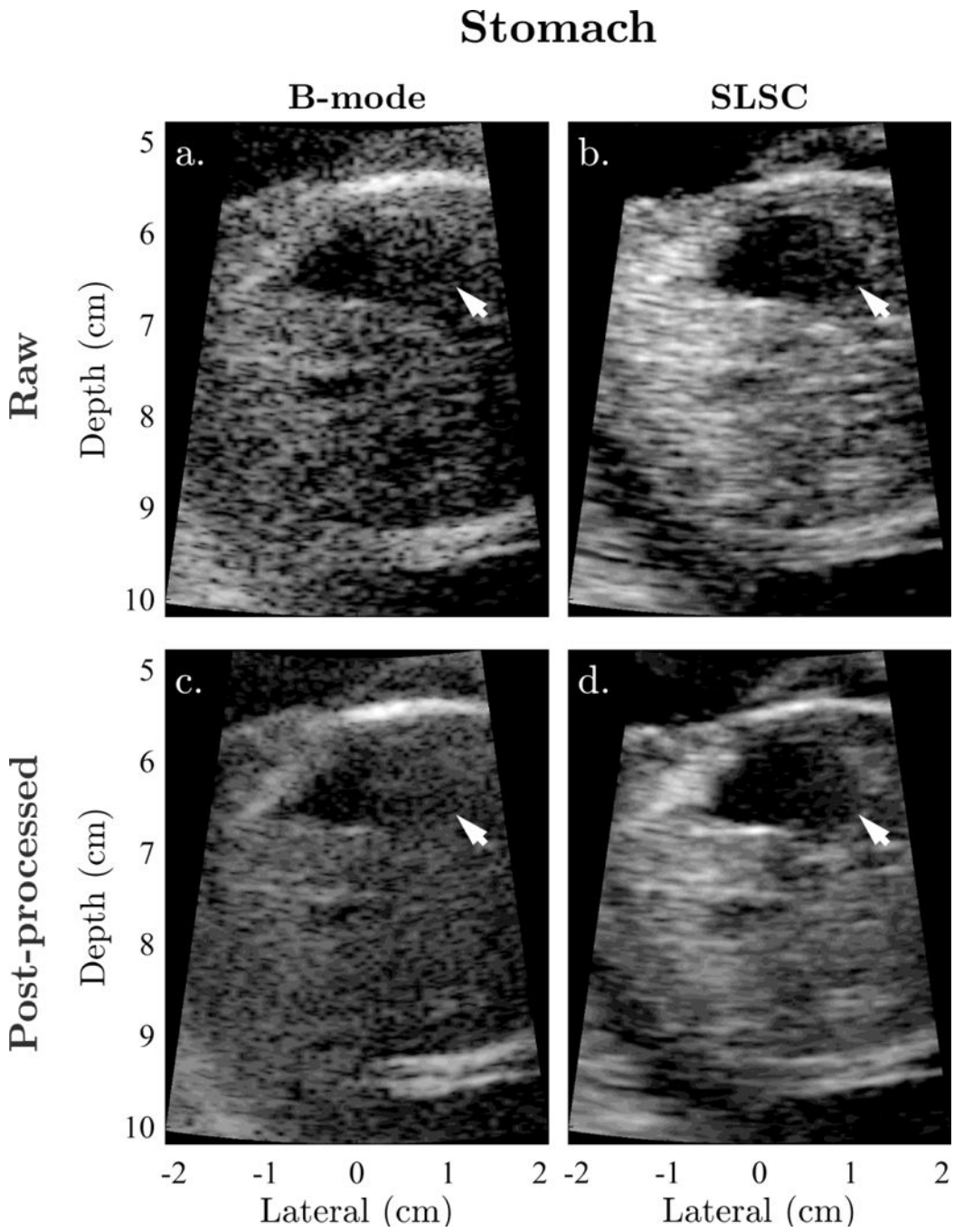


Figure 2.

Example matched (a) raw B-mode, (b) raw SLSC, (c) post-processed B-mode, and (d) post-processed SLSC images of the fetal stomach shown with sonographer-optimized parameters for dynamic range and short-lag. Arrows indicate the lateral boundary of the stomach, which is significantly degraded by reverberation clutter in B-mode. SLSC imaging appears to reduce the appearance of this clutter, resulting in improved conspicuity and definition of the stomach and its boundaries.

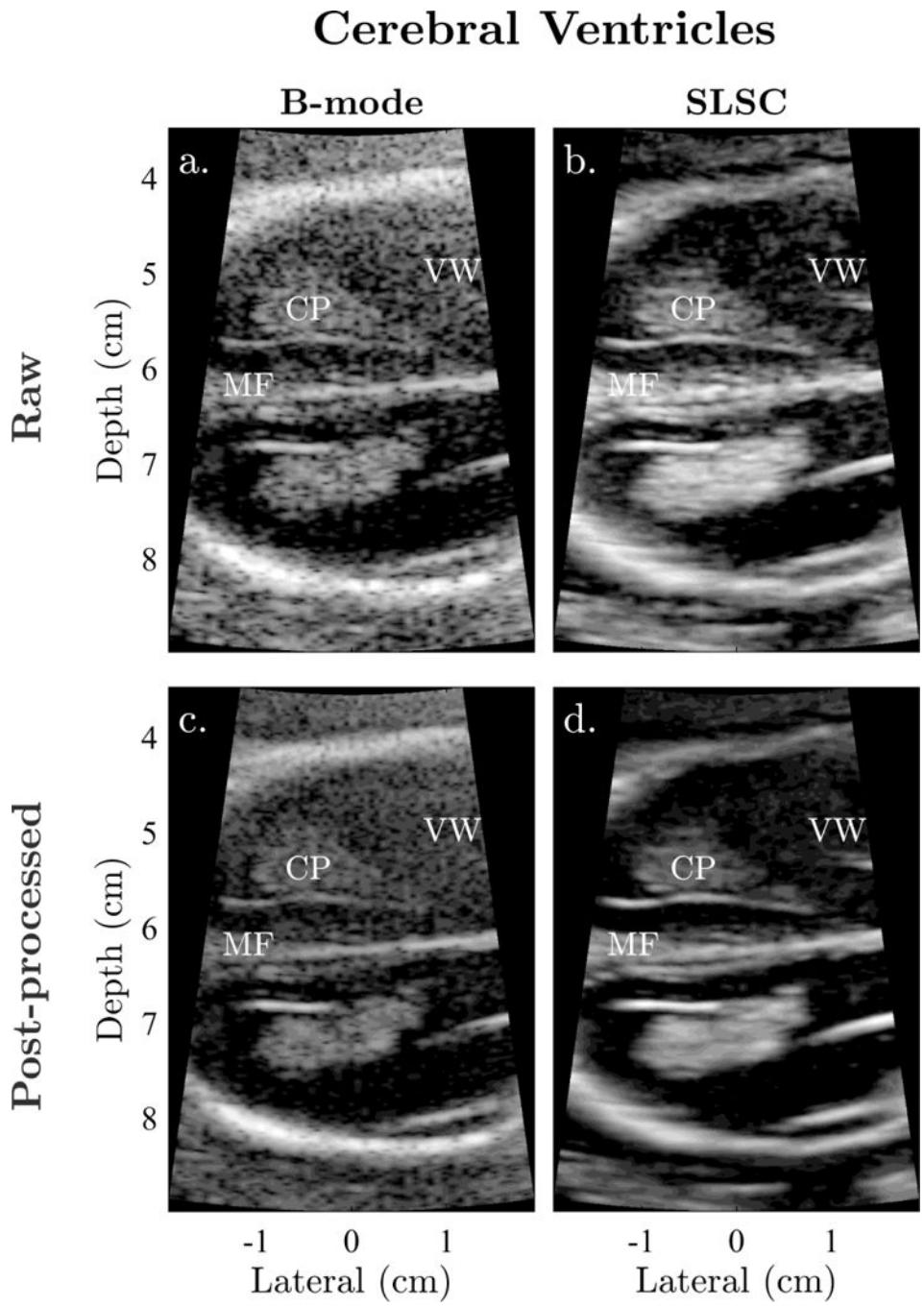


Figure 3.

Example matched (a) raw B-mode, (b) raw SLSC, (c) post-processed B-mode, and (d) post-processed SLSC images of the fetal cerebral ventricles shown with sonographer-optimized parameters for dynamic range and short-lag. In both raw and post-processed B-mode images, reverberation from the uterine wall and off-axis clutter result in poor visibility of the choroid plexus (CP) and ventricular walls (VW). These structures appear better visualized in SLSC as a result of clutter suppression.

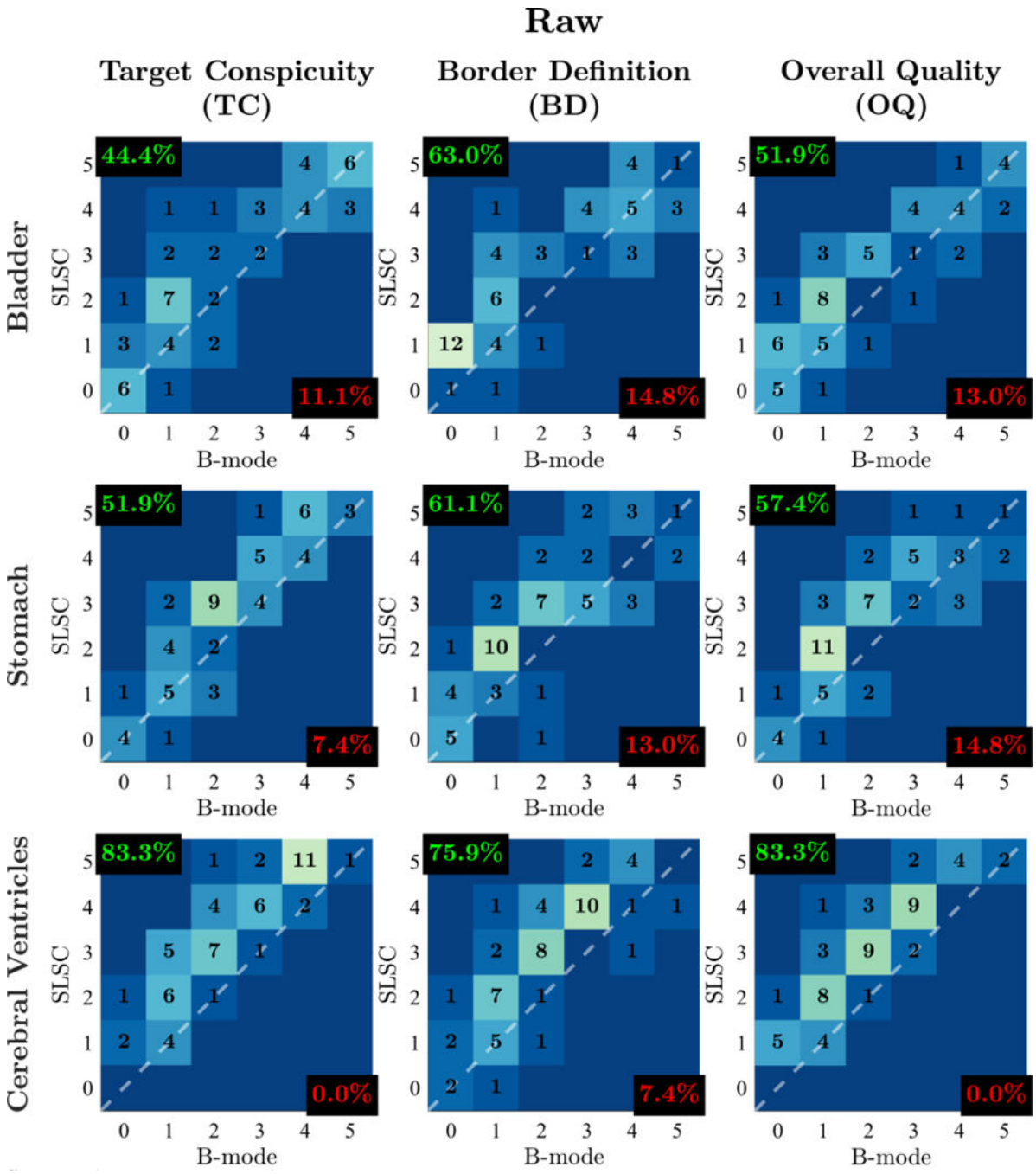


Figure 4.

Outcome measure score distributions for raw video pairs of the fetal bladder, stomach, and cerebral ventricles. Histograms are organized in columns for different outcome measures (TC, BD, OQ) and in rows for different imaged structures. Individual histograms depict the number of matched video pairs in different score categories with SLSC score plotted as a function of B-mode score. The dashed white lines denote categories in which SLSC and B-mode scores are the same. Video pairs above this diagonal indicate increase in outcome measure score from B-mode to SLSC. For each histogram, the percentage of video pairs

with SLSC preference and B-mode preference are indicated in the top-left and bottom-right corners, respectively.

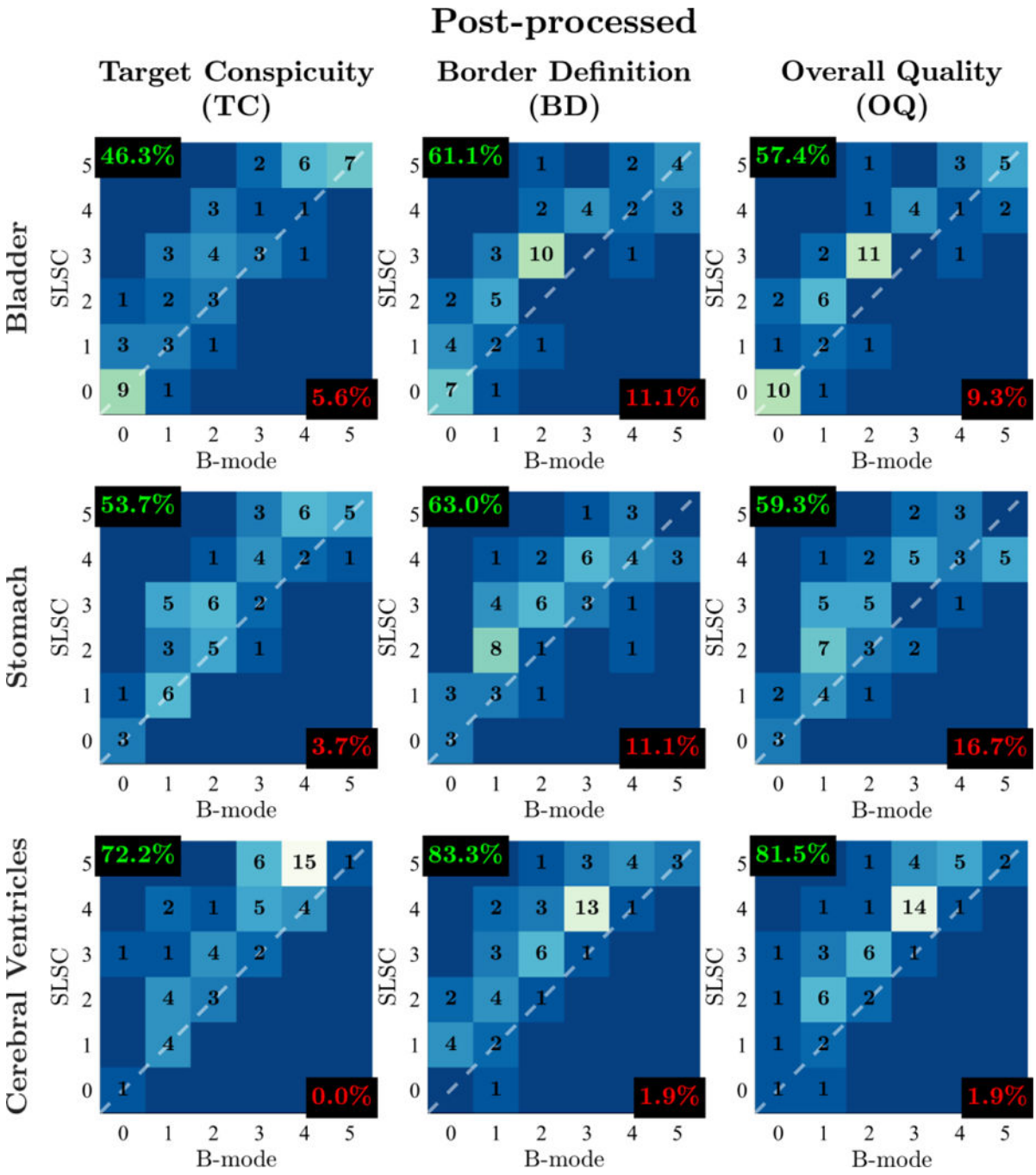


Figure 5.

Outcome measure score distributions for post-processed video pairs of the fetal bladder, stomach, and cerebral ventricles. Histograms are organized in columns for different outcome measures (TC, BD, OQ) and in rows for different imaged structures. Individual histograms depict the number of matched video pairs in different score categories with SLSC score plotted as a function of B-mode score. The dashed white lines denote categories in which SLSC and B-mode scores are the same. Video pairs above this diagonal indicate increase in outcome measure score from B-mode to SLSC. For each histogram, the percentage of video

pairs with SLSC preference and B-mode preference are indicated in the top-left and bottom-right corners, respectively.

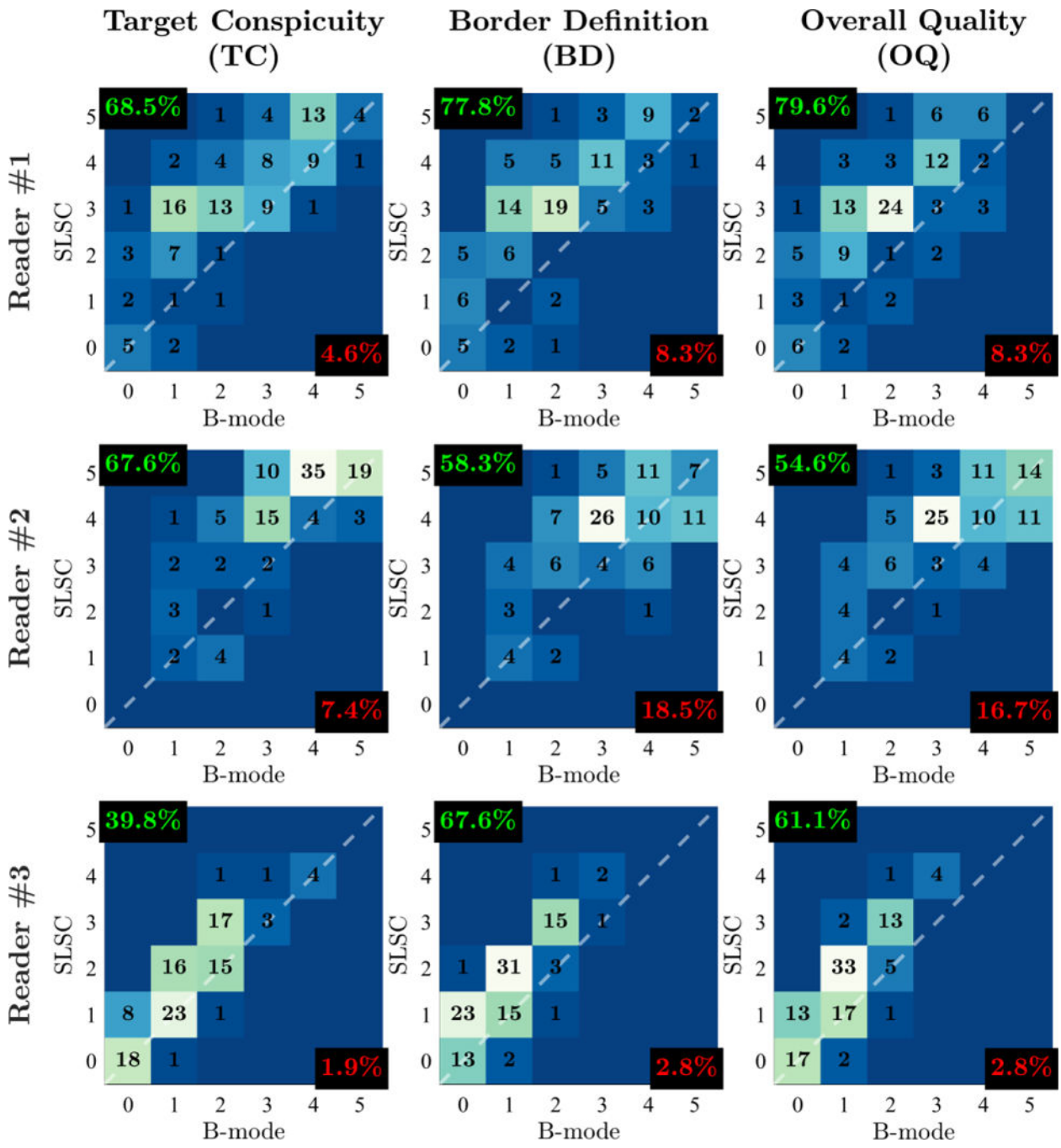


Figure 6.

Outcome measure score distributions for Readers 1, 2, and 3. Histograms are organized in columns for different outcome measures (TC, BD, OQ) and in rows for different readers. Individual histograms depict the number of matched video pairs in different score categories with SLSC score plotted as a function of B-mode score. The dashed white lines denote categories in which SLSC and B-mode scores are the same. Video pairs above this diagonal indicate increase in outcome measure score from B-mode to SLSC. For each histogram, the

percentage of video pairs with SLSC preference and B-mode preference are indicated in the top-left and bottom-right corners, respectively.

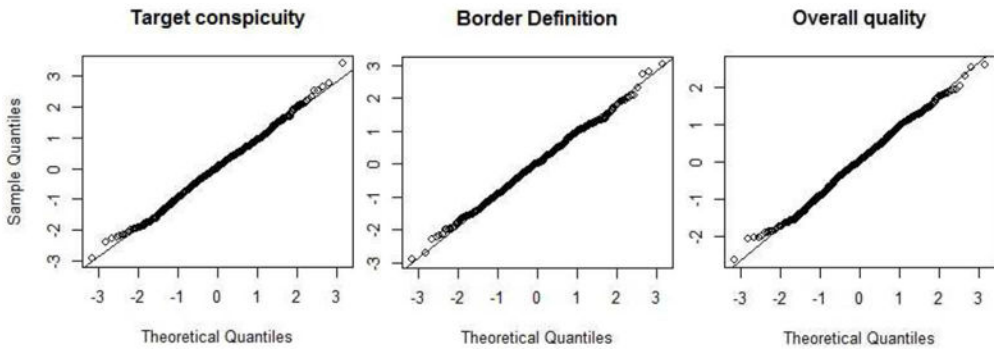
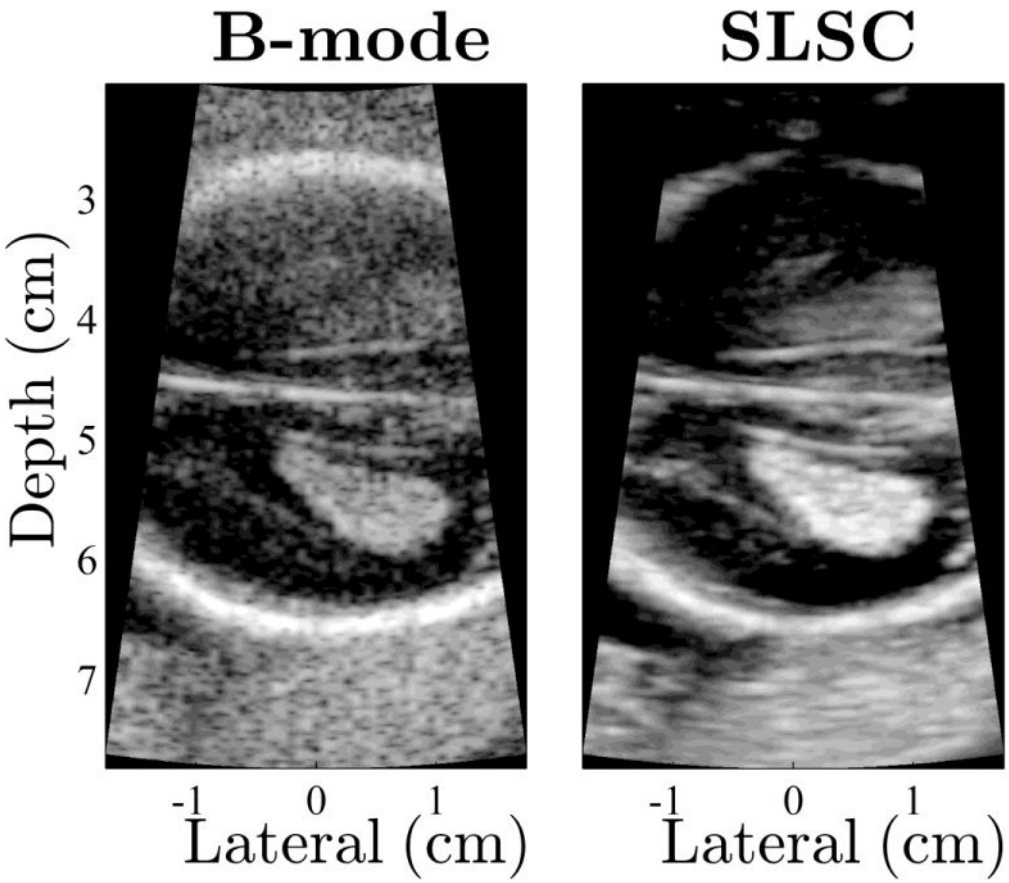


Figure 7.

Quantiles of estimated errors from the multivariate model fit plotted against the quantiles of the standard normal distribution for all three outcome measures. Linearity of the points in each plot indicates strong adherence of model fit residuals to the Gaussian distribution.



Video Figure 1.

Example video showing matched B-mode and SLSC image sequences of the cerebral ventricles generated from raw channel data acquired using the modified clinical system. Image sequences are able to successfully capture fetal and probe motion at 14 fps over the 3-second acquisition.

Table 1

Reported BMI values.

Body mass index (kg/m^2)	# of subjects
<25	3
25–30	10
31–40	5
>40	0

Table 2

Acoustic output measurements.

	Measured Output	FDA Limit	Units
MI	1.62	≤ 1.9	–
$I_{SPTA.3}$	54	≤ 720	mW/cm^2
$I_{SPPA.3}$	160	≤ 190	W/cm^2

MI = mechanical index

$I_{SPTA.3}$ = derated spatial-peak temporal-average intensity

$I_{SPPA.3}$ = derated spatial-peak pulse-average intensity

Table 3

Image optimization and scoring criteria.

Criterion	Description
Target Conspicuity (TC)	How well can you discern the target from the background?
Border Definition (BD)	How sharp are the edges of the target?
Overall Quality (OQ)	What is the quality of the image as a whole?

Table 4

Multivariate model with interactions for target conspicuity (TC) scores.

Term	Effect	Standard Error	*t* value	*p*-value
Baseline	2.13	0.74	2.9	0
SLSC	0.47	0.15	3.04	0.002*
Post-processing	0.14	0.11	1.25	0.21
Stomach	0.17	0.13	1.26	0.21
Cerebral Ventricles	0.32	0.13	2.44	0.01*
BMI	−0.01	0.04	−0.31	0.76
SLSC:Post-processing	0.06	0.15	0.4	0.69
SLSC:Stomach	0.07	0.19	0.39	0.69
SLSC:Cerebral Ventricles	0.53	0.19	2.81	0.005*
SLSC:BMI	0.03	0.02	1.5	0.13

* $p < 0.05$

Table 5

Multivariate model with interactions for border definition (BD) scores.

Term	Effect	Standard Error	*t* value	*p*-value
Baseline	1.91	0.69	2.78	0.01
SLSC	0.55	0.15	3.77	0.0002 *
Post-processing	0.15	0.1	1.44	0.15
Stomach	0.06	0.13	0.51	0.61
Cerebral Ventricles	0.15	0.13	1.17	0.24
BMI	−0.03	0.04	−0.69	0.49
SLSC:Post-processing	0.13	0.15	0.89	0.37
SLSC:Stomach	0	0.18	0	1
SLSC:Cerebral Ventricles	0.38	0.18	2.13	0.03 *
SLSC:BMI	0.03	0.02	2.13	0.03 *

* $p < 0.05$

Table 6

Multivariate model with interactions for overall quality (OQ) scores.

Term	Effect	Standard Error	*t* value	*p*-value
Baseline	1.88	0.69	2.71	0.01
SLSC	0.48	0.14	3.38	0.0007*
Post-processing	0.17	0.1	1.73	0.08
Stomach	0.16	0.12	1.29	0.2
Cerebral Ventricles	0.19	0.12	1.51	0.13
BMI	−0.02	0.04	−0.66	0.51
SLSC:Post-processing	0.09	0.14	0.66	0.51
SLSC:Stomach	0.05	0.17	0.27	0.79
SLSC:Cerebral Ventricles	0.52	0.17	3	0.003*
SLSC:BMI	0.03	0.02	1.79	0.07

* $p < 0.05$

Table 7

Inter-rater agreement as measured by Spearman's correlation.

Outcome Measure	Readers		
	1 and 2	2 and 3	1 and 3
Target Conspicuity (TC)	0.71	0.58	0.48
Border Definition (BD)	0.72	0.62	0.59
Overall Quality (OQ)	0.75	0.66	0.58